AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE

10.1002/2014JD021953

Key Points:

- Multiple satellite data sets are combined to satisfy the water budget closure
- In situ data comparison demonstrates effectiveness of closure postfiltering
- A Closure Correction Model allows to calibrate satellite data sets independently

Correspondence to:
S. Munier,
simon.munier@gmail.com

Citation:
Munier, S., F. Aires, S. Schlaffer, C. Prigent, F. Papa, P. Maisongrande, and M. Pan (2014), Combining data sets of satellite-retrieved products for basin-scale water balance study: 2. Evaluation on the Mississippi Basin and closure correction model, *J. Geophys. Res. Atmos.*, *119*, 12,100–12,116, doi:10.1002/2014JD021953.

Received 1 MAY 2014
Accepted 19 OCT 2014
Accepted article online 22 OCT 2014
Published online 13 NOV 2014

Combining data sets of satellite-retrieved products for basin-scale water balance study: 2. Evaluation on the Mississippi Basin and closure correction model

Simon Munier[1], Filipe Aires[1,2], Stefan Schlaffer[3], Catherine Prigent[2], Fabrice Papa[4,5], Philippe Maisongrande[4], and Ming Pan[6]

[1]Estellus/LERMA, Observatoire de Paris, Paris, France, [2]Laboratoire de l'Etude du Rayonnement et de la Matière en Astrophysique, CNRS, Observatoire de Paris, Paris, France, [3]Department of Geodesy and Geoinformation, Research Group Remote Sensing, Vienna University of Technology, Vienna, Austria, [4]Laboratoire d'Etude en Géophysique et Océanographie Spatiales-CNES-CNRS-IRD-UPS, Toulouse, France, [5]Indo-French Cell for Water Sciences, IISc-NIO-IITM-IRD Joint International Laboratory, Indian Institute of Science, Bangalore, India, [6]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA

Abstract In this study, we applied the integration methodology developed in the companion paper by Aires (2014) by using real satellite observations over the Mississippi Basin. The methodology provides basin-scale estimates of the four water budget components (precipitation P, evapotranspiration E, water storage change ΔS, and runoff R) in a two-step process: the Simple Weighting (SW) integration and a Postprocessing Filtering (PF) that imposes the water budget closure. A comparison with in situ observations of P and E demonstrated that PF improved the estimation of both components. A Closure Correction Model (CCM) has been derived from the integrated product (SW+PF) that allows to correct each observation data set independently, unlike the SW+PF method which requires simultaneous estimates of the four components. The CCM allows to standardize the various data sets for each component and highly decrease the budget residual ($P - E - \Delta S - R$). As a direct application, the CCM was combined with the water budget equation to reconstruct missing values in any component. Results of a Monte Carlo experiment with synthetic gaps demonstrated the good performances of the method, except for the runoff data that has a variability of the same order of magnitude as the budget residual. Similarly, we proposed a reconstruction of ΔS between 1990 and 2002 where no Gravity Recovery and Climate Experiment data are available. Unlike most of the studies dealing with the water budget closure at the basin scale, only satellite observations and in situ runoff measurements are used. Consequently, the integrated data sets are model independent and can be used for model calibration or validation.

1. Introduction

Under a changing climate, the global hydrological cycle is expected to accelerate and intensify [e.g., *Del Genio et al.*, 1991; *Trenberth*, 1999; *Huntington*, 2006; *Coumou and Rahmstorf*, 2012]. In regions not limited by water availability, the Clausius-Clapeyron relation implies that the specific humidity should increase with temperature, leading to modifications of the different components of the water cycle, including precipitation, evapotranspiration, river discharge, and freshwater stored on continents. Significant societal impacts are very likely, with, for instance, effects on water availability for the population or modifications of the drought and flood patterns. Changes have already been observed on precipitation [e.g., *Dai et al.*, 2004] or evapotranspiration [*Jung et al.*, 2010]. *Labat et al.* [2004] reported a probable intensification of the global water cycle with an increasing global runoff (total discharge from the continents to the oceans), but more recent studies [*Dai et al.*, 2009; *Milliman et al.*, 2008; *Munier et al.*, 2012] did not observe any significant trend on the global runoff over the last several decades, showing that no consensus has been reached on that topic in the scientific community.

One of the main reasons for the lack of consensus is that consistent description of the present and past hydrological cycle is still not available at the global scale with the needed accuracy, despite significant efforts at international level, within the Global Energy and Water Experiment (GEWEX) for instance. Various global hydrological data sets have been developed, some of them from satellites, for the different water cycle components such as precipitation (e.g., Global Precipitation Climatology Project (GPCP)), evapotranspiration (e.g., Global Land surface Evaporation: the Amsterdam Methodology (GLEAM)), continental water storage (e.g., Gravity Recovery and Climate Experiment (GRACE)), or river discharges (e.g., Global Runoff Data Centre (GRDC)). At this time, no large-scale satellite data set on river discharges exists, even though some efforts are made to derive the discharge in large rivers from remote sensing data, namely in preparation for the forthcoming Surface Water and Ocean Topography satellite mission [e.g., *Bjerklie et al.*, 2005; *Gleason and Smith*, 2014; *Durand et al.*, 2014]. For some of these variables, different global data sets exist, but they are showing significant differences in time and spatial variability with strong difficulties to evaluate the accuracy of each individual data set. In addition, consistency between these different data sets has been questioned. At the basin scale, the different hydrological variables should at least close the water budget; i.e., the water budget equation should be satisfied:

$$\frac{dS}{dt} = P - E - R \tag{1}$$

where P is the precipitation, E the evapotranspiration, R the runoff (or the discharge at the outlet of the basin), and S the total water storage (sum of water stored in the vegetation, snow, lakes and rivers, soil moisture, and groundwater). All the components are given in millimeters. Besides, the runoff (streamflow) is converted from m^3/s to millimeters by dividing it by the basin area.

Several recent studies have investigated the water budget closure from the key hydrological observations in equation (1). Most of them pointed out the difficulty to close this budget due to the insufficient accuracy of the individual data sets and to their inconsistencies [e.g., *Azarderakhsh et al.*, 2011]. As a consequence, experiments have been conducted to close this budget [*Sheffield et al.*, 2009; *Sahoo et al.*, 2011; *Pan et al.*, 2012] by using an assimilation strategy within a land surface model. For instance, following a Kalman filter algorithm with an additional constraint related to the budget closure, *Pan et al.* [2012] combined multiple data sets and model outputs to extract a coherent data set of the four components (P, E, R, and S) over different basins around the world.

In a companion paper [*Aires*, 2014], several methodologies were proposed and tested, to ensure a consistent integration of the different hydrological data sets. With a synthetic experiment, it was shown that the water budget can be closed without any assimilation techniques and without using any model. This is an important point since one of the potential applications of the integrated global products is to calibrate or validate land surface or climate models.

In this study, one of the integration technique presented in the companion paper [*Aires*, 2014] is applied to satellite-derived hydrological products and tested over a well-documented basin. The most widely used global data sets for precipitation, evapotranspiration, and continental water storage are integrated with runoff data derived from in situ measurements. The analysis focuses on the Mississippi Basin, which is an important basin in terms of hydrological processes and socioeconomic impact [*Milly and Dunne*, 2001]. In addition, many in situ measurements are available over this basin, allowing us to evaluate the results of our integration process. A closure correction method is further introduced that transforms independently each satellite data set toward an integrated data set that satisfies the budget closure.

The Mississippi Basin is first presented in section 2, together with a description of the data sets used in this study. The integration methodology is introduced in section 3. The results are presented in section 4, with an evaluation of the integrated data set with in situ data from rain gauges and from the AmeriFlux experiment. Section 5 presents the closure correction method obtained from results of section 4. Gap filling applications are also presented, with a comparison to products previously obtained from assimilation techniques and model outputs [*Pan et al.*, 2012]. Finally, section 6 concludes this study and gives some possible perspectives.

2. Case Study and Data Sets

2.1. Mississippi Basin

The Mississippi Basin is located in central USA (Figure 1). It is one of the best-observed continental-scale basins in the world. In addition to the satellite observations that are used in this study, model-based estimates and in situ observations of the various budget components are also available for comparison. Streamflow is relatively accurately monitored using in situ gauge measurements [*Sheffield et al.*, 2009]. *Milly and Dunne* [2001] showed that runoff increased by 22% from 1949 to 1997 and precipitation by 10% (which suggests a probable intensification of the water cycle). Over a more recent period (1993–2007), the surface

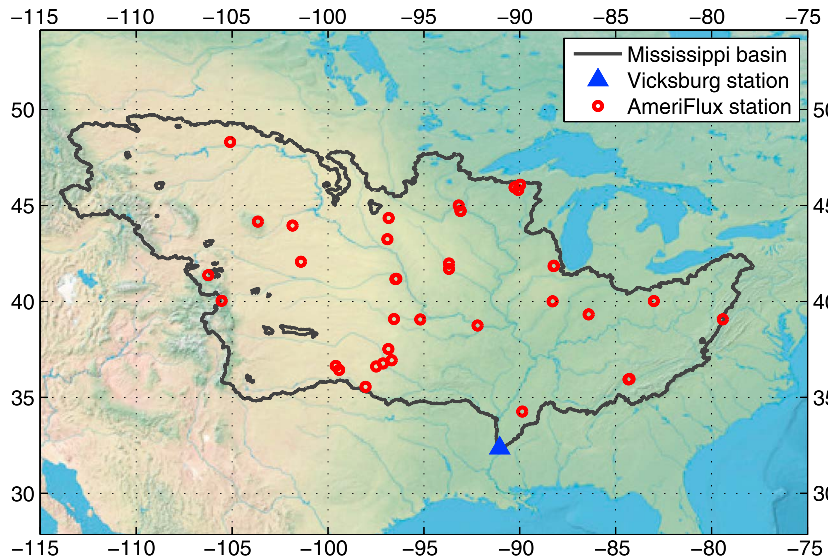

Figure 1. Mississippi Basin upstream of the Vicksburg station and location of AmeriFlux stations.

water extent significantly decreased by 48.5% [*Prigent et al.*, 2012]. In addition, the water cycle in the Mississippi Basin is also affected by human alterations, e.g., abandonment of cultivated farmlands and conversion to forest or pasture as well as large impoundments [*Milly and Dunne*, 2001].

In this study we considered the basin upstream the Vicksburg gauging station, with an area of 2,964,255 km^2. The basin delineation was provided by the Global Runoff Data Centre (GRDC) and aggregated to a 0.5° spatial resolution. The mean discharge at the Vicksburg station is 17,600 m^3/s with a significant part due to snowmelt.

2.2. Data Sets

This section describes the data sets used in this study. Some of them are based on satellite products and are used in the integration process. The others are based on ground observations and are used only for validation purposes. Since no satellite-based runoff product with a sufficient temporal resolution exists, in situ runoff (streamflow) observations are used in the integration process. The characteristics of the different data sets are summarized in Table 1.

2.2.1. Data Sets Used in the Integration Process

Except for the runoff data which was extracted from the Global Runoff Data Centre (GRDC, Vicksburg streamflow gauge station, monthly observations), all the other data sets are derived from satellite observations.

We considered four precipitation data sets: the Tropical Rainfall Measuring Mission (TRMM, 3B43 V7) Multisatellite Precipitation Analysis (TMPA), the NOAA CPC Morphing Technique (CMORPH, V1.0), the Naval Research Laboratory (NRL) blended technique, and the Global Precipitation Climatology Project (GPCP, V2.2). It has to be noticed that the TMPA and GPCP products have been corrected using in situ rain gauge observations. Several studies were dedicated to the validation of these data sets over different basins with various hydroclimatic conditions such as *Huffman et al.* [2007], *Su et al.* [2008], *Villarini et al.* [2009], *Scheel et al.* [2011], *Wu et al.* [2014] for TMPA, *Joyce et al.* [2004], *Kidd et al.* [2011], *Sheffield et al.* [2009] for CMORPH, *Kidd et al.* [2011] for NRL, and *Gosset et al.* [2013] for GPCP. Overall, these products compare well with rain gauge observations at the monthly time scale, even though large biases can affect daily rainfall amount estimates. Other studies focused on biases and uncertainties of these products [*Sapiano and Arkin*, 2009; *Tian and Peters-Lidard*, 2010; *Thiemig et al.*, 2012; *Gosset et al.*, 2013]. Although these studies provided quite precise quantifications of uncertainties, they also showed that uncertainties clearly depend on the basin and the meteorological conditions (wet/dry season, complex terrain, convective precipitation, etc.). Unfortunately, no study provides a comprehensive uncertainty analysis of the four products over the Mississippi Basin.

Table 1. Data Sources for the Four Components and Main Characteristics[a]

Name	Source	Period	Spatial Resolution	Reference
		Precipitation (*P*)		
CRU	Rain gauges	1901–present	0.5°	*Mitchell and Jones* [2005]
WM	Rain gauges	1900–present	0.5°	*Willmott and Matsuura* [2010]
GPCC	Rain gauges	1900–present	0.5°	*Schneider et al.* [2008]
TMPA	Satellite	1998–present	0.25°	*Huffman et al.* [2007]
CMORPH	Satellite	1998–present	0.25°	*Joyce et al.* [2004]
NRL	Satellite	2003–2010	0.25°	*Turk et al.* [2010]
GPCP	Satellite	1979–present	2.5°	*Adler et al.* [2003]
		Evapotranspiration (*E*)		
GLEAM	Satellite	1980–2011	0.25°	*Miralles et al.* [2011]
MOD16	Satellite	2000–2012	1 km	*Mu et al.* [2007]
NTSG	Satellite	1983–2006	8 km	*Zhang et al.* [2010]
		Water Storage Change (ΔS)		
CSR	Satellite	2002–present	Basin	http://grace.jpl.nasa.gov/data/
GFZ	Satellite	2002–present	Basin	http://grace.jpl.nasa.gov/data/
JPL	Satellite	2002–present	Basin	http://grace.jpl.nasa.gov/data/
GRGS	Satellite	2002–present	Basin	http://grgs.obs-mip.fr/grace/
		River Discharge (*R*)		
GRDC	Gauges	1900–present	Basin	http://www.grdc.sr.unh.edu/

[a]The three gauge-based precipitation products (Climate Research Unit (CRU), Willmott and Matsuura (WM), and Global Precipitation Climatology Centre (GPCC)) are used only for validation purposes. TMPA, Tropical Rainfall Measuring Mission Multisatellite Precipitation Analysis; CMORPH, Climate Prediction Center (CPC) Morphing Technique; NRL, Naval Research Laboratory; NTSG, Numerical Terradynamic Simulation Group; MOD16, Moderate Resolution Imaging Spectroradiometer (MODIS) Global Evapotranspiration Project; CSR, Center for Space Research; GFZ, German Research Centre for Geosciences; JPL, Jet Propulsion Laboratory; GRGS, Groupe de Recherche en Géodésie Spatiale.

For the evapotranspiration, three products were chosen: Global Land surface Evaporation: the Amsterdam Methodology (GLEAM), MODIS Global Evapotranspiration Project (MOD16), and NTSG Land Surface Evapotranspiration (NTSG). Validation of each data set against eddy flux tower observations can be found in *Miralles et al.* [2011], *Mu et al.* [2011], and *Zhang et al.* [2010], respectively, showing a good comparison at basin scale and at monthly time scale. Satellite-derived evapotranspiration (*E*) is the result of modeling actual *E* using energy-balance schemes such as the Penman-Monteith (NTSG and MOD16 products) or the Priestly and Taylor (GLEAM) approaches. Therefore, uncertainties are mainly a combination of the errors of the meteorological input data and the errors introduced by the model. According to *Ferguson et al.* [2010], the biggest source of uncertainty is the vegetation parameterization, followed by surface temperature and radiation.

The continental water storage variations were estimated using four products, all of them based on the Gravity Recovery and Climate Experiment (GRACE, *Tapley et al.* [2004]) but obtained with different preprocessing and postprocessing: Jet Propulsion Laboratory (JPL), Center for Space Research (CSR), German Research Centre for Geosciences (GFZ), and Groupe de Recherche en Géodésie Spatiale (GRGS). Numerous studies can be found in the literature on the validation of GRACE data for hydrological applications [e.g., *Klees et al.*, 2008; *Syed et al.*, 2008; *Werth and Güntner*, 2010; *Forootan et al.*, 2012]. The main limitation of GRACE data is its spatial resolution (about 300 km) which limits its use to large river basins. The GRACE-Tellus monthly grids include both measurement and leakage errors which have been described by *Landerer and Swenson* [2012]. In addition to the measurement error, the leakage error describes the error introduced by the weighted spatial averaging of the GRACE data when converting from spectral (spherical harmonics) to geographical coordinates. Due to the strong spatial correlation of the errors, they are reduced when averaging storage change over large basins.

Each of the data sets described in this section has been interpolated at a 0.5° spatial resolution, spatially averaged over the Mississippi Basin and aggregated to monthly values as this was found to be the lowest temporal resolution among all data sets (especially ΔS). All the data sets are presented in Figure 2.

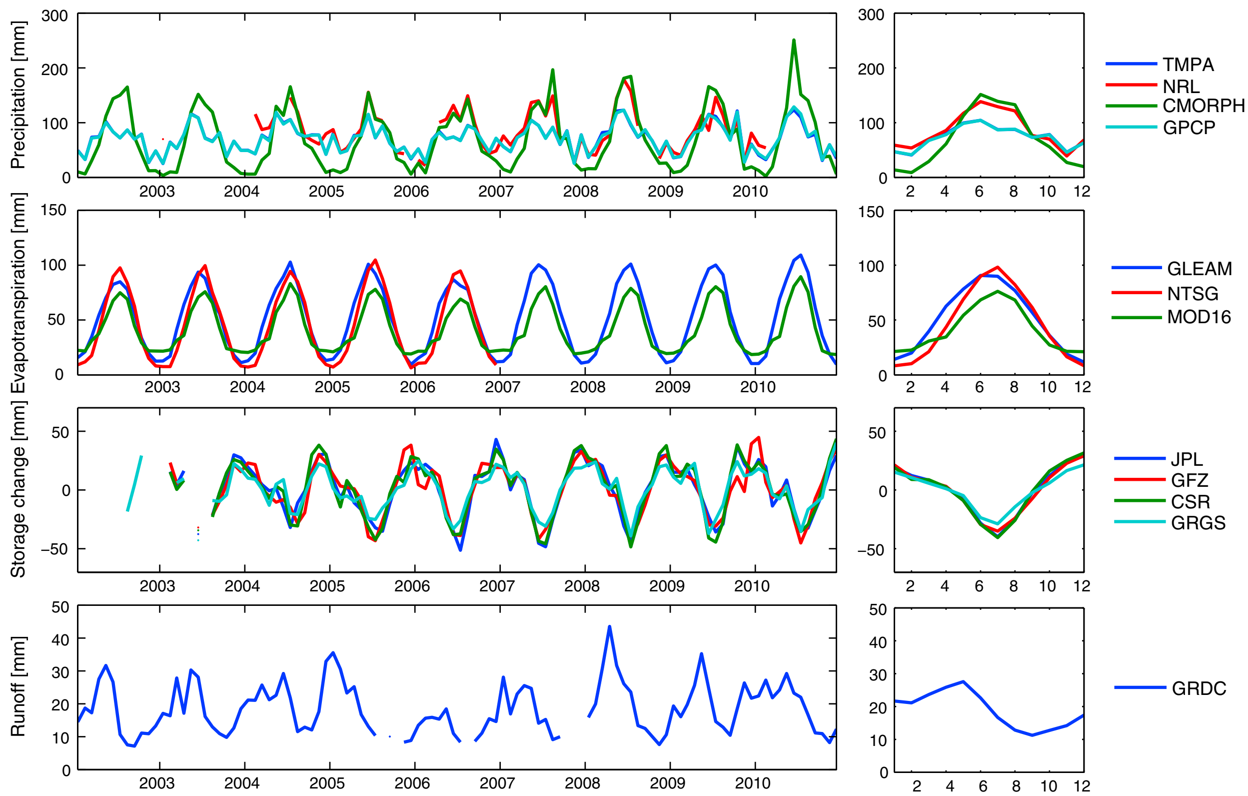

Figure 2. Observations of the four components of the water budget used in the integration process (time series on the left column and seasonal means on the right column). The precipitation, evapotranspiration, and storage change data sets are satellite based, while the runoff is ground based (Vicksburg streamflow gauge station).

Note that time lags due to flow propagation inside the basin are accounted for in the ΔS component. For instance, if a rainfall event occurs in the upstream part of the basin, the water storage will increase and remain large until the water reaches the outlet (or evaporates). Until then, the runoff will start to increase, and the water storage will start to decrease.

2.2.2. Data Sets Used for Validation

Other data sets only based on rain gauge measurements are used in this study for validation purposes. Namely, we considered the following global-gridded precipitation data sets: Global Precipitation Climatology Centre (GPCC), Climate Research Unit (CRU), and Willmott and Matsuura (WM).

We also used observations from in situ stations from AmeriFlux for precipitation and evapotranspiration (see Figure 1 for their locations). The AmeriFlux network coordinates regional analysis of observations from eddy covariance flux towers across North America, Central America, and South America [*Law et al.*, 2006]. We obtained the Level 4 data product for 39 AmeriFlux sites over the period 2000–2006 from the AmeriFlux website (http://public.ornl.gov/ameriflux).

2.3. Uncertainties

Some studies aimed at characterizing the uncertainties of satellite-retrieved products (see section 2.2.1). Nevertheless, such characterizations are generally product and site specific, and for some products used in this study, no uncertainty characterization can be found in the literature, or at least not for the Mississippi Basin. For that reason we considered the same uncertainty for all the data sets of a given parameter, i.e., a white noise with a standard deviation of 10 mm/month for P and E, 5 mm/month for ΔS, and 1 mm/month for R. The choice of these values was motivated by results of the studies cited in section 2.2.1 for the P [*Sapiano and Arkin*, 2009; *Thiemig et al.*, 2012; *Gosset et al.*, 2013] and E [*Ferguson et al.*, 2010] data sets. For example, *Gosset et al.* [2013] and *Sapiano and Arkin* [2009] estimated an error of about 10% for TMPA, which corresponds approximately to 10 mm/month if we assume that there is no seasonality in this uncertainty. For the GLEAM product, *Miralles et al.* [2011] estimated an error of about 17%. The FLUXNET stations used in their study shows a mean annual evapotranspiration of 560 mm. Hence, assuming the uncertainty to be constant throughout the year, the value of 10 mm/month is also realistic. For ΔS, this value seems reasonable considering the great size of the Mississippi Basin and values given in other studies [*Landerer et al.*, 2010; *Syed et al.*, 2005; *Zaitchik et al.*, 2008]. Finally, it was assumed in this study that the runoff uncertainties was at most 5% of the mean discharge, which is about 1 mm/month.

3. Method

3.1. Bias Correction

Figure 2 shows that large systematic discrepancies exist between the different estimates of a water component, especially for P and E. However, as specified in the first part paper [*Aires*, 2014], the statistical methodology developed here, as most of the statistical methods, is based on the assumption that observations are unbiased. The difficulty is that, as for uncertainties, it is rather difficult to obtain bias estimates from the literature for every data set used in this study. Without any a priori information on biases, the average state of the components used to compute biases is chosen to the average of all the databases. Then, all the sources of information for this component are bias corrected toward this mean.

3.2. Integration Method

In this section, the notations and the integration methodology are briefly recounted, and more details are provided in the companion paper [*Aires*, 2014]. We first consider the four water components $X^T = (P, E, R, \Delta S)$ (^{T} is the transpose symbol). The closure of the water budget is obtained when $X^T \cdot G = 0$, where $G^T = (1, -1, -1, -1)$, which is equivalent to the water budget equation (equation (1)).

Let:

$$Y_\epsilon{}^T = (P_1, \dots, P_p, E_1, \dots, E_q, R_1, \dots, R_r, \Delta S_1, \dots, \Delta S_s) \tag{2}$$

be the vector of dimension $n = p + q + r + s$ gathering the multiple observations available for each water component: (1) $(P_1, P_2, \dots, P_p)$, the p precipitation estimates; (2) $(E_1, E_2, \dots, E_q)$, the q evapotranspiration estimates; (3) $(R_1, R_2, \dots, R_r)$, the r runoff estimates; and (4) $(\Delta S_1, \Delta S_2, \dots, \Delta Ss)$, the s water storage change estimates.

The aim of this study is to obtain a linear filter K used to obtain an estimate X_a ("a" stands for analysis) of X based on the observations Y_ϵ:

$$X_a^{\mathrm{sw}} = K \cdot Y_\epsilon, \tag{3}$$

where K is a $4 \times n$ matrix. Several methods were considered in the companion paper [*Aires*, 2014], but Monte Carlo simulations have shown that for this particular application (i.e., the closure of the terrestrial water cycle), the so-called Simple Weighting (SW) plus Postprocessing Filtering (PF) procedure provides results as good as more complex techniques. In the SW approach, each water component is weighted based on its a priori uncertainties (see equation (8) in *Aires* [2014] and section 2.3).

The SW filter does not impose any closure constraint on X_a. An interesting Postprocessing Filtering approach has been introduced, like, e.g., in *Pan and Wood* [2006], to impose the closure constraint on a previously obtained solution $X_b = K \cdot Y_\epsilon$. This strategy was used in a Kalman filtering context in *Pan and Wood* [2006], using a first guess estimate with some constraints from the Variable Infiltration Capacity (VIC) [*Liang et al.*, 1994, 1996] macroscale hydrologic model. In *Aires* [2014], the PF was used and tested without any model, as a simple postprocessing step after the SW:

$$X_a^{\mathrm{pf}} = K_{\mathrm{pf}} \cdot X_b, \tag{4}$$

where $K_{\mathrm{pf}} = (\mathrm{Id} - BG(G^T BG)^{-1} G^T)$, Id is the identity matrix, and B is the error covariance matrix of the first estimate X_b.

At a monthly time scale, evapotranspiration mainly driven by vegetation growth and can be considered as a low-frequency signal [*Allen et al.*, 1998]. Since the postfiltering step (budget closure) consists in partitioning the budget residual among the four components at each time step independently, high frequencies may appear in the resulting estimation of E. To avoid this problem, we enforced the budget closure by frequencies. We first decomposed each parameter into a high- and a low-frequency components. To choose a relevant cutoff frequency, we first performed a fast Fourier transform (FFT) decomposition of the evapotranspiration data sets. The decomposition clearly showed substantially smaller signals for frequencies

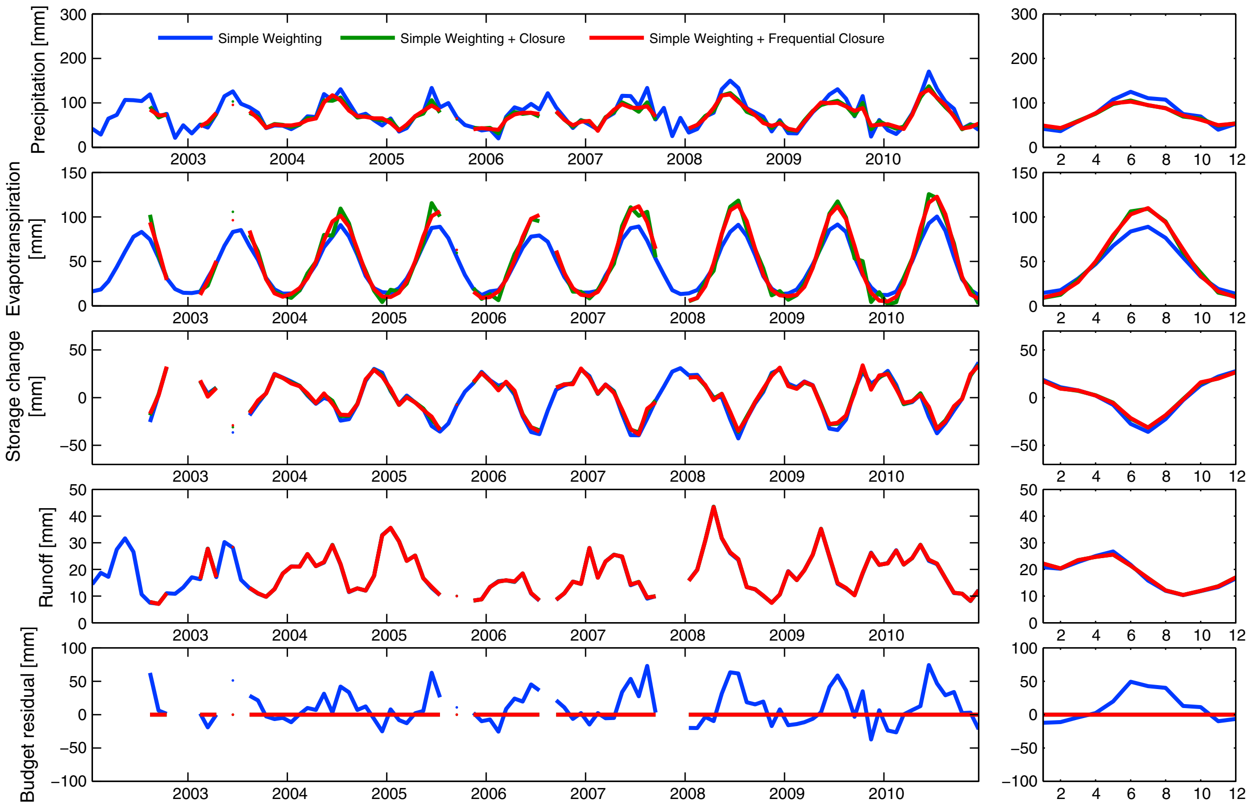

Figure 3. (left column) Time series and (right column) seasonal means of the four components and the budget residual obtained with the Simple Weighting integration with and without the simple/frequential budget closure postfiltering.

greater than 6 months. Thus, a cutoff frequency of 6 months was chosen to separate high and low frequencies of each component (using an FFT decomposition). The budget closure is then applied on high and low frequencies independently. The two resulting data sets are combined to obtain the integrated product. The high-frequency component of E is supposed to be null and is then not included in the high-frequency budget closure. The linearity of PF and frequency transformations ensures the budget closure of the final product.

4. Integration Results

In this section, the Simple Weighting plus Postfiltering integration technique is applied on observation data sets described in section 2.2. The resulting integrated product is then analyzed and validated against in situ observations.

4.1. Simple Weighting and Budget Closure

Figure 3 shows the results of the integration and the budget closure steps for the four water cycle components. The bottom plot represents the water budget residual, i.e., the quantity $P - E - R - \Delta S$.

Since the uncertainties are chosen equal for the various source of information of each water component, the SW integration computes the average of the available data sets for each component. If no closure constraint is imposed, the water budget residual ranges from −40 to 80 mm/month. The seasonal budget residual shows an excess of water (positive residual) during the wet season (from May to October) and a shortfall of water (negative residual) during the dry season (from November to March).

We then applied the PF without the frequency decomposition. The PF step is able to distribute the excesses and shortfalls of water among the four components depending on their relative uncertainties. Since the highest uncertainties have been assigned to the precipitation and evapotranspiration, they are logically the most modified water components [*Pan and Wood*, 2010]. On the contrary, the observed discharge was

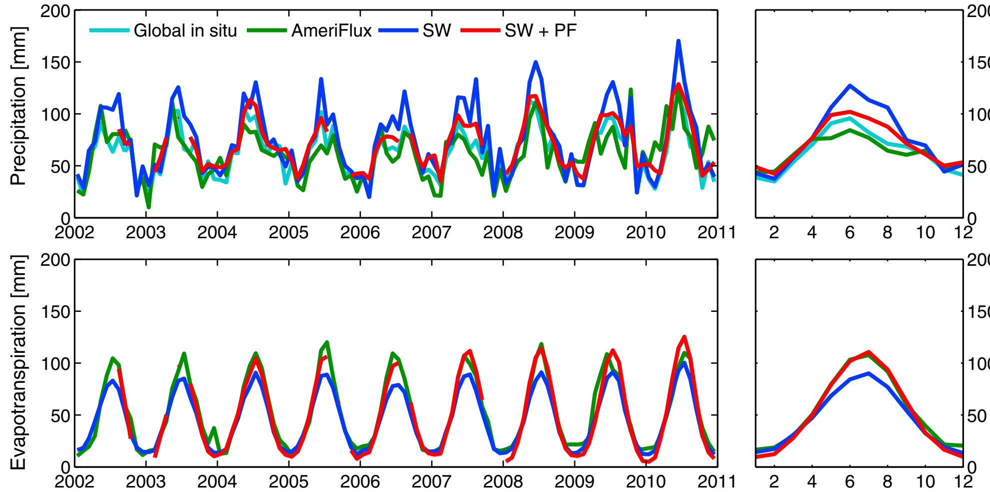

Figure 4. Comparison of the integrated product (with and without the postfiltering process) to in situ precipitation and evapotranspiration. Data sources are from rain gauge-based global data sets (for precipitation only) and from the AmeriFlux data set (for both precipitation and evapotranspiration).

assumed to be the most reliable; therefore, it is only slightly modified by the PF process. The four resulting components are plotted in green in Figure 3. As expected, the PF step decreases the precipitation and increases the evapotranspiration during the wet season. The storage change is also slightly increased, and the runoff shows no significant change. As expected, the budget is closed at each time step.

Note that the evapotranspiration obtained after the PF step shows some high-frequency signals in particular during the wettest and the driest months. These high-frequency signals in the integrated components are directly related to the high frequencies of the water budget residual (lower graph, in blue). Applying the PF by frequencies highly reduced these high-frequency signals in E, which is shown by the red line in Figure 3. As for the regular SW+PF integration (without the frequency decomposition), its frequency-decomposed version obtains the closure of the water budget since the frequency decomposition and recomposition, as well as the PF process, are linear. In the following, the frequency decomposed PF is always chosen.

All the four components are needed to apply the budget closure postfiltering (equation (4)). If any component is missing, the other components cannot be updated, resulting in additional gaps in the integrated product (see, e.g., missing values in ΔS in 2003 or in R in 2005 and 2008).

Finally, the postfiltering process has little impact on the overall variability of the components. Indeed, the correlation between the integrated data sets before and after the postfiltering reaches 0.98 or more for the four components. Results are similar regarding the interannual variability, except for the evapotranspiration for which the correlation is only 0.59. Applying the frequency-decomposed postfiltering increases this value to 0.82 but simultaneously decreases the precipitation interannual correlation from 0.93 to 0.76. The correlations for ΔS and R remain unchanged, greater than 0.98. This is an important positive argument for the SW+PF integration method. This approach is able to introduce coherency among the water components by imposing the budget closure, with a minimal impact on the temporal variability of the original data sets.

4.2. Evaluation Using In Situ Data

The integrated products are compared to in situ observations for evaluation purposes. Two sources of in situ measurements are used (see section 2.2.2): (1) global-gridded data sets based on rain gauge data are used to evaluate precipitation and (2) the AmeriFlux database also serves as a reference for both precipitation and evapotranspiration. For the AmeriFlux data, spatial averages of P and E have been estimated over the Mississippi Basin using the Thiessen polygons weighting method: information from each rain gauge is weighted by the area of the Thiessen polygon associated to the rain gauge location. The comparison results are presented in Figure 4.

Some differences exist between the gridded and the AmeriFlux precipitation data sets. They may be due to different interpolation techniques and additional rain gauge stations accounted for in the global data sets.

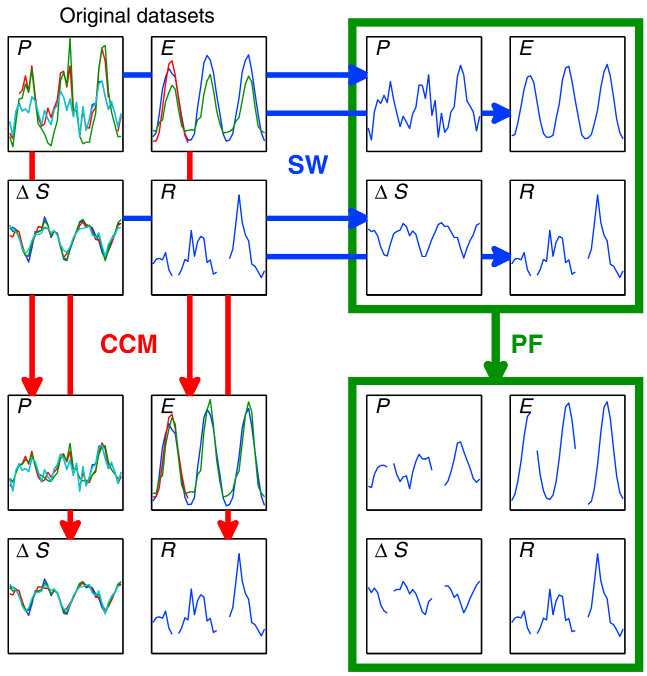

Figure 5. Integration and postfiltering processes (SW+PF) and closure correction model (CCM) methodology. SW stands for Simple Weighting and PF for Postprocessing Filtering. The CCM is calibrated against the SW+PF results and applies on each original data set independently.

At this stage, it is difficult to state that one data set is better than the other, and we preferred to consider both for the comparison. The correlation is 0.79, and root-mean-square error (RMSE) is about 15 mm/month.

Compared to the in situ products, the SW integration overestimates the precipitation and underestimates the evapotranspiration during the wet summer. This is partly corrected when applying the PF process, which decreases the RMSE (with respect to the AmeriFlux data sets) for the summer months (May to September) from 27.3 to 12.4 mm/month for P and from 16.0 to 8.9 mm/month for E. The PF process is slightly less efficient during winters, leading to a slightly reduced improvement when considering the whole time period: the RMSE decreases from 18.6 to 11.8 mm/month for P and from 11.4 to 8.8 mm/month for E. According to *Sapiano and Arkin* [2009], the bias in CMORPH and NRL is small in winter but large in summer, which may explain the smallest improvement in winter. In every case, the correlation between the in situ observations and the integrated product (with or without postfiltering) ranges from 0.81 to 0.98. As stated in the previous section, the postfiltering does not change the correlation significantly.

5. Closure Correction Model and Applications

5.1. Closure Correction Model

To obtain the integrated set of components, multiple data sets are combined, independently, for each water component. But the closure constraint of the PF step combines also together each one of the four water component estimates. For instance, the integration procedure with PF corrects precipitations from TMPA using other precipitation data sets but also estimates of E, ΔS, and R. In this section, a "closure correction model" (CCM) is introduced to correct each data set independently, based on the results of the SW+PF integration. For instance, a closure correction model specific to TMPA would directly correct the TMPA precipitation, without the use of other components (E, Q, and ΔS) and without the use of the other precipitation data sets. Figure 5 illustrates the CCM methodology. As shown in the figure, the CCM, which is calibrated using results of the SW+PF integration, allows to correct each data set independently. The resulting product (CCM+SW) is obtained without the use of the other components, unlike the SW+PF integration process.

The closure correction model is represented by simple affine transformations defined by a scaling factor a and an offset b such that $X_{\mathrm{new}} = a \cdot X + b$. The model parameters a and b are calibrated by computing linear regressions between the original observation data sets and the SW+PF integrated components. The method allows the bias correction (see section 3.1) to be included in the closure correction model. Results of the linear regressions are shown in Figure 6, while values of the regression parameters a and b are presented in Table 2.

First, the linear regression is quite relevant since the r^2 statistics (i.e., percentage of variance of the target data explained by the model) is higher than 0.75 for all the data sets. Such a result shows that each data set can be corrected independently with a satisfactory accuracy.

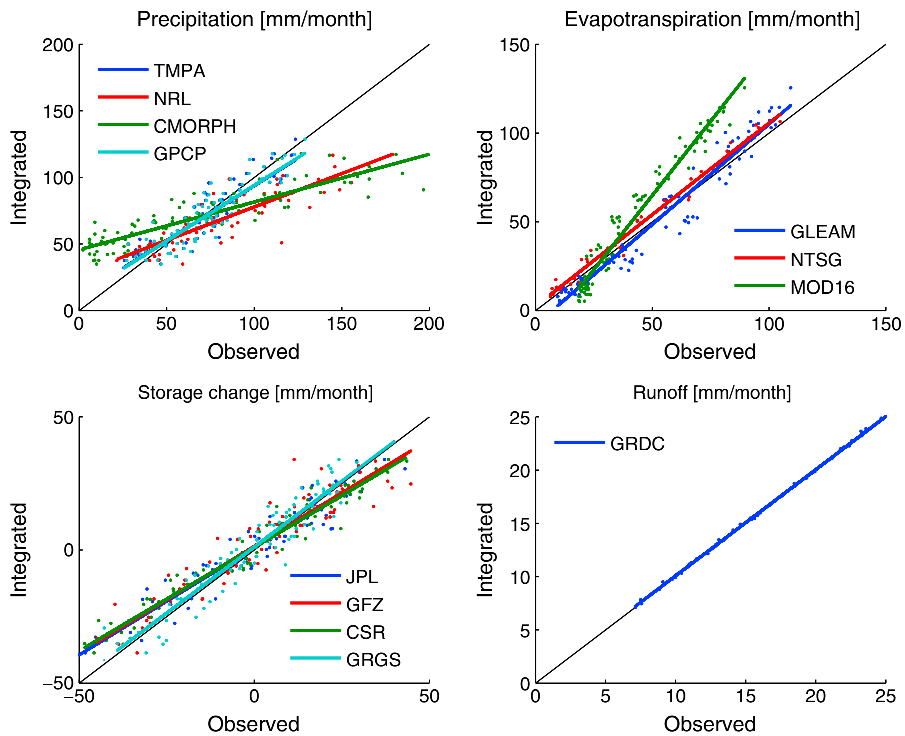

Figure 6. Linear regressions for the closure correction model for the four components data sets used in the integration process.

Concerning the precipitation component, TMPA and GPCP observations are very close to the integrated product (*a* close to 1 and low values of *b*). This result is consistent with the fact that in situ observations are used to calibrate both satellite data sets, even though this property was not used in the integration process. This figure also shows that NRL and CMORPH overestimate the amplitude of the seasonal cycle, which confirms results from *Sapiano and Arkin* [2009]. Yet the closure correction model is able to rescale them with a satisfactory accuracy (r^2 reaches 0.75 and 0.81, respectively).

Because of the strong seasonality of evapotranspiration (and very low interannual variability), the closure correction model is very efficient, with r^2 correlation coefficients higher than 0.95 for the three data sets. The GLEAM and NTSG products are close to the integrated product, unlike MOD16 that clearly underestimates E (high value of a).

The regression coefficients for the three official releases of GRACE (i.e., JPL, GFZ, and CSR) are very close to each other since their preprocessing and postprocessing is quite similar. All of them slightly overestimate the storage change compared to the integrated component. On the other hand, the GRGS product differs in its processing and seems to be more accurate in this case ($a = 0.99$). The linear regression approximation is quite efficient for this component too, with r^2 statistics ranging from 0.89 to 0.95.

Finally, the closure correction model has almost no impact on the runoff. This is due to the low uncertainty attributed to this component during the integration process. Even though in situ discharge measurement

Table 2. Regression Coefficients *a* and *b* of the Closure Correction Model[a]

	TMPA	NRL	CMORPH	GPCP	GLEAM	NTSG	MOD16	JPL	GFZ	CSR	GRGS	GRDC
a	0.82	0.50	0.36	0.83	1.13	1.03	1.67	0.81	0.80	0.78	0.99	1.00
b	11.29	27.70	45.67	10.57	−8.03	2.37	−18.72	0.97	1.33	0.99	1.09	0.11
r^2	0.80	0.75	0.81	0.81	0.95	0.99	0.96	0.93	0.89	0.95	0.90	1.00

[a]The r^2 statistics is also indicated as a quality indicator of the regressions.

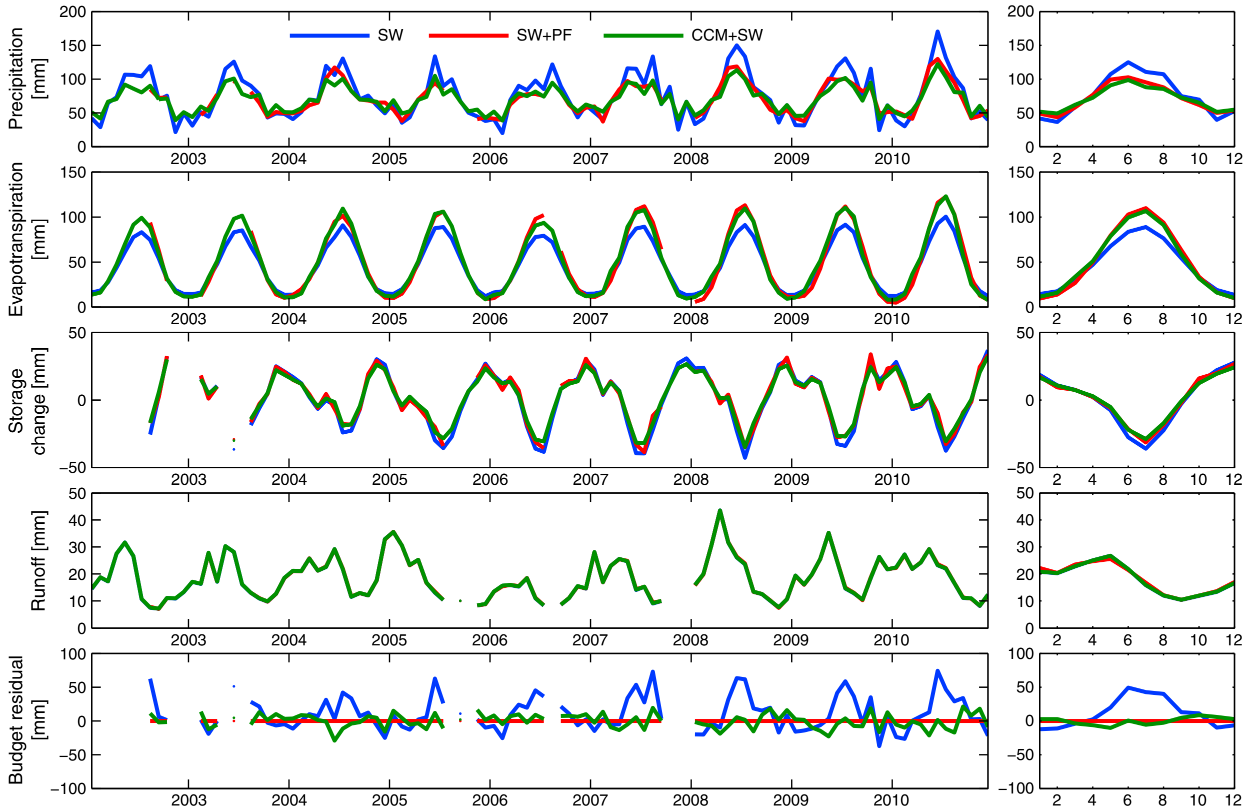

Figure 7. (left column) Times series and (right column) seasonal means of the integrated product (with and without the closure postfiltering) and the closure correction model results. The budget residual is also shown.

may suffer from uncertainties, it was assumed in this study that such uncertainties were at most 5% of the mean discharge, which is 1 mm/month. This value being much lower than uncertainties in the other components; the discharge is assumed to be much more reliable and is consequently only slightly modified during the integration process.

The high r^2 values given in Table 2 show that the CCM can be accurately approximated using a simple linear regression. It implies that only a few years may be required to calibrate the CCM, provided that available data cover a sufficiently wide range of possible values. It is quite difficult to give a minimum length of model training period, because it strongly depends on the range of available data. Basically, considering only 1 year with a typical hydrological cycle (e.g., 2005) could give the same results as considering the whole time period. Nevertheless, a longer time period would provide more robust results. Another factor that could influence the reliability of the CCM is the correlation between the original data and the integrated product. Indeed, the CCM is a simple linear transformation and does not impact the temporal variability and the correlation. Consequently, if the SW+PF process has little impact on the temporal variability (which depends on the original data sets and on the studied basin), the CCM would be more reliable.

5.2. Evaluation of the Closure Correction Model

As stated in the previous section, applying the CCM to the observation data sets makes each of them closer to the integrated product. The mean standard deviation between all the data sets decreased from 15.9 to 5.5 mm/month for P, from 7.2 to 5.5 mm/month for E, and from 5.0 to 3.8 mm/month for ΔS.

To evaluate the quality of the CCM, three integrated products are compared and shown in Figure 7. The first one, the SW integration, corresponds to the simplest approximation obtained by averaging all the available observation data sets, independently for each component. The second one, the SW+PF integration, corresponds to our best estimate obtained by combining all the data sets with a budget closure constraint. The third one is obtained with the SW integration after having applied the CCM on each data set independently.

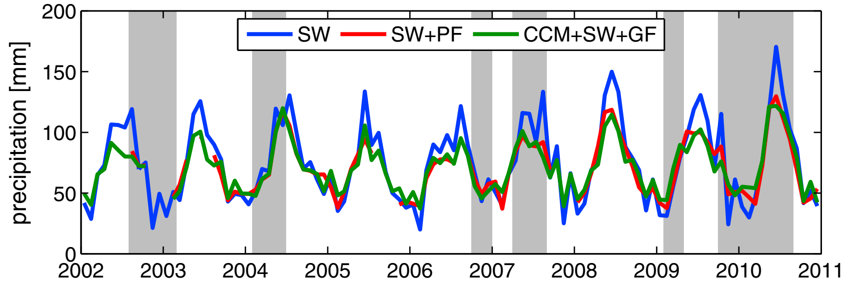

Figure 8. Example of precipitation gap filling using the closure correction model. GP stands for gap filling. Grey shadings represent simulated gaps. The SW and SW+PF curves are obtained without the simulated gaps and shown as references.

Applying the CCM before the integration process greatly influences the results (see the comparison between the blue and green curves). Moreover, the impact of the CCM is highly positive since the resulting components are very close to the SW+PF results. The RMSE between CCM and SW+PF decreased from 14.8 to 8.0 mm/month for P, from 10.6 to 4.0 mm/month for E, and from 3.7 to 2.9 mm/month for ΔS when applying the CCM prior to the SW integration. The CCM step not only improves each component estimate but it is also beneficial to the water budget closure: the water budget residual ($P - E - R - \Delta S$) is significantly reduced, with a standard deviation decreasing from 25 mm/month for SW to 10 mm/month for CCM+SW.

5.3. Application: Gap Filling

The PF process is not able to compute any of the four components when one of them is missing. For instance, in Figure 7, some discharge values are missing at the end of year 2007, leading to missing values in P, E, and ΔS for SW+PF, even though there is no gap in those three components in the original observation data sets. Since the closure correction model is applied for each component independently, the computation of the three components is therefore possible. This important advantage allows for an interesting application for time series gap filling. Basically, when one component is missing, the three other components are corrected using the CCM approach, and the missing component is computed from the water budget closure equation (equation (1)). This gap filling approach will be referred to CCM+closure in the following.

To evaluate this strategy, gaps in the precipitation data sets have been artificially introduced in the observation data sets. P is then reconstructed whenever it is possible (i.e., when none of the three other components is missing). An example is shown in Figure 8. Blue and red curves represent the SW and SW+PF estimates, respectively, if no data were missing. The SW+PF red curve is the best estimate of P, and the reconstructed missing values should be as close as possible to it. Simulated gaps are represented by grey shadings. The green curve represents the CCM results outside the grey shadings (same as Figure 7), but inside the grey shadings, it represents the gap filling from the CCM+closure approach. The gap filling method is clearly able to reconstruct the precipitation, even when a whole year of data is missing (e.g., in 2010). Note that gaps remain when any of the three other components is missing (e.g., end of 2002).

To have a clearer idea of the ability of the method to reconstruct missing values, a Monte Carlo experiment has been set up to obtain some gap filling statistics. As previously mentioned, random gaps have been generated to simulate missing value periods, and the gap-filled values obtained with CCM+closure are compared to the artificially suppressed data. To better evaluate the results of this experiment, the CCM+closure gap filling results are also compared to two other simple gap filling approaches. The first one uses a linear interpolation to fill in the missing values of the water components. The second gap filling method is based on a “seasonal linear regression”: a linear regression between the available observations and the seasonal mean is performed and used to fill in the gaps.

For the three methods, the PF process is applied after the gap filling step, to impose the closure constraint to the gap-filled data sets. The experiment was run 1000 times for each component (a stability analysis has been performed and showed that this number was large enough to provide robust results). For the three methods, we computed the average RMSE between the reconstructed values and the integrated product obtained if no data were missing. Note that only the reconstructed values are used to compute the RMSE (not the entire period). Table 3 collects the results.

Table 3. Gap Filling Monte Carlo Experiment: Average RMSE (in mm/Month) Between Target and Gap-Filled Estimates, for the Three Filling Methods: Linear Interpolation of the Missing Component, Seasonal Regression, and Closure Correction Model (CCM) Approach

	P	*E*	ΔS	*R*
Linear interpolation	13.96	19.04	15.11	6.87
Seasonal regression	8.42	1.43	9.15	7.87
CM+closure	5.41	5.01	8.94	10.85

Except for the runoff, the linear interpolation shows the highest RMSE values, as expected. Better results are obtained with the seasonal regression. The RMSE for *E* is particularly low (1.43 mm/month), which highlights the high seasonality of the evapotranspiration signal. The gap filling is improved with the CCM+closure approach for the *P* and ΔS components. For *E*, the CCM+closure gap filling is less efficient than the seasonal regression. This is due to the strong seasonality of the evapotranspiration which is well captured by the seasonal regression method.

Results for the reconstruction of the runoff *R* are very different. The highest RMSE is actually obtained with the CCM+closure gap filling method. Indeed, in the case of the Mississippi Basin, the standard deviation of the runoff is only 7.7 mm/month, whereas it is 22.7, 36.4, and 18.8 mm/month for *P*, *E*, and ΔS, respectively. In comparison, the standard deviation of the budget residual is 9.3 mm/month for CCM+SW (it is 25.1 mm/month for SW). In the gap filling methodology, we assume that the budget is closed so that the budget equation can be applied to compute any component from the three others. The fact that the runoff variations have the same order of magnitude as the budget residuals explains the limited performances of the runoff gap filling with the closure correction model (CCM+closure). A simple linear interpolation of the runoff gives more accurate results than the combination of the other water components which are more uncertain.

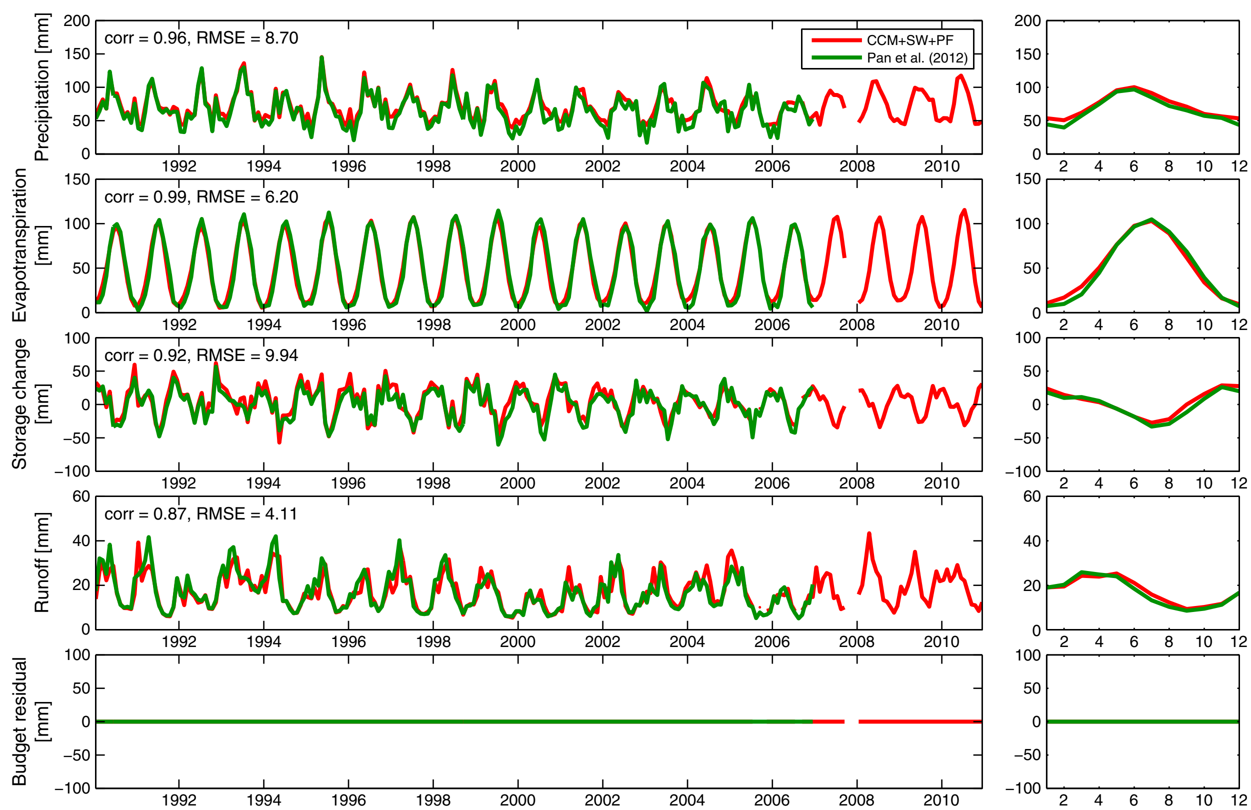

Figure 9. Reconstruction of past water storage change using CCM and comparison with results from *Pan et al.* [2012]. (left column) Time series of the four components and the budget residual are presented, as well as (right column) seasonal means. The correlation (corr) and root-mean-square error (RMSE) between both results are indicated in the upper left corner for each component.

5.4. Application: Reconstruction of Past Ground Water Storage Change

GRACE observations are not available before 2002. However, the CCM+PF gap filling technique can potentially be used to estimate the ground water storage change. Observations of P, E, and R are available over the period 1990–2002, so their estimates before 2002 can be used to estimate ΔS. Figure 9 compares such a reconstruction with the one from the *Pan et al.* [2012] study. The latter was obtained using a constrained data assimilation technique ensuring the water budget closure [*Pan and Wood*, 2006]. The VIC model was used in combination with GRACE observations to estimate ΔS before and after 2002. The three observed components (P, E, and R) are consistent with results from *Pan et al.* [2012], with RMSE ranging from 4.1 to 8.7 mm/month and correlation ranging from 0.87 to 0.99. Concerning the storage change, the comparison also gives very good results (RMSE of 9.9 mm/month and correlation of 0.92), as well as for the interannual signal (RMSE of 9.9 mm/month and correlation of 0.77).

6. Discussion and Conclusion

In this study, we applied the integration methodology developed in *Aires* [2014] by using real satellite observations over the Mississippi Basin. The methodology provides estimates of the four water budget components: precipitation P, evapotranspiration E, water storage change ΔS, and runoff R. This integration is a two-step process. The first step is the Simple Weighting (SW) integration. With the assumption of identical uncertainties for all the data sets of each component, SW is equivalent to a simple average. The second step is a Postprocessing Filtering (PF) process that imposes, at the basin level, the water budget closure by distributing the budget residual ($P - E - \Delta S - R$) among the four components, depending on their respective uncertainties. In this study, the PF process described in *Aires* [2014] has been improved by decomposing each component into low and high frequencies. This avoided spurious high-frequency signals in E that appear when redistributing the budget residual among the four components. The integrated product (SW+PF) consists in basin-scale monthly time series of the four components.

A comparison with in situ observations of P and E, namely from the AmeriFlux database and various rain gauge-based global data sets, showed the benefits of the PF process. Indeed, the precipitation derived by SW was overestimated during the summer, while the evapotranspiration was underestimated. Differences with in situ observations were highly reduced when the PF step was applied, showing that the closure constraint can be used to improve our characterization of the water cycle.

The GPCP and TMPA precipitation products, which are used in the integration process, are corrected with rain gauge measurements, and logically compare better to in situ observations than the NRL and CMORPH data sets. Nevertheless, it is quite difficult to assess the quality of these data sets in ungauged regions where they may not provide the best estimates of P. In addition, the in situ rain gauge network may not be dense enough in some regions (e.g., Northwestern mountains of the Mississippi Basin) and may be biased by a limited spatial representativeness. Hence, in order to provide a methodology as general as possible, we preferred to include the NRL and CMORPH data sets into the integration process. Interestingly, results showed that the SW+PF process was able to satisfactorily correct them.

To compute the integrated product, all the four water cycle components are needed. This means that if one of them is missing, which happens frequently for satellite data sets, the methodology cannot be applied directly. A closure correction model (CCM) has been derived to correct each observation data set independently, without the use of the other observations. This model is based on a linear regression that transforms the original observations into more coherent observations. It has been calibrated using the results of SW+PF integration to obtain solutions closer to the budget closure. Results of the CCM compare well with the SW+PF integrated products; CCM can then be considered as equivalent to SW+PF. The water budget computed with the CCM product is not closed, but the residuals are highly reduced and small enough for most applications. The CCM is a satellite data set calibration procedure that facilitates the integration of multiple data sets and takes into account the water budget closure. It can be noted that this closure model is very simple since it is based on a linear regression of the original data sets (but calibrated on the SW+PF results). Figure 10 shows a synthesis of the overall methodology for the precipitation component. The upper part presents the SW+PF strategy. Note that the four components are needed to apply the PF step. Alternatively, the CCM, which is calibrated against results from SW+PF, is presented on the lower part. The figure clearly shows that the CCM can be applied on each data set independently and pixelwise. The resulting

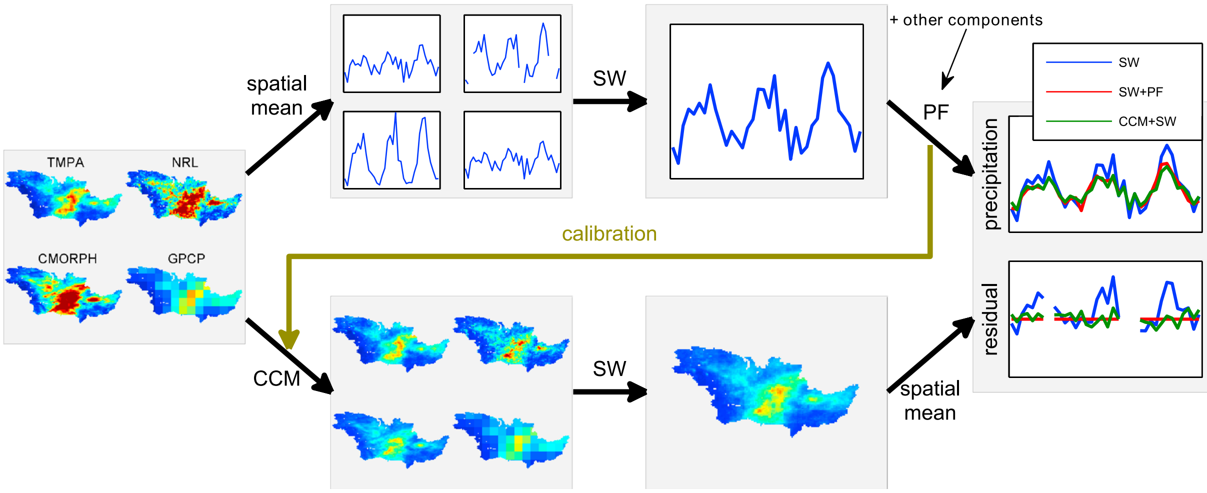

Figure 10. Methodology synthesis: example of the precipitation component. Integration and postfiltering processes (SW+PF, upper part) and closure correction model (CCM, lower part) methodology. The four components are required to apply the PF process. The CCM is calibrated against the SW+PF results and can be applied on each original data set independently and pixelwise.

product compares well to the SW+PF product, and the budget residual is highly reduced compared to the SW product (without the closure constraint provided by PF).

When the four components are available, the SW+PF should be preferred to the CCM, at the basin scale. However, one of the main advantages of the closure correction model is that it is possible to estimate the water components and ensure the closure of the water budget, even when one of these components is missing. A direct application is the use of the CCM to fill gaps in one of the components: the CCM transformation is applied to the three available components and the closure equation allows to estimate the missing fourth one. Results of a Monte Carlo experiment with synthetic gaps demonstrated the good performances of this CCM+closure methodology, except for the runoff data that has a variability of the same order of magnitude as the budget residual. The discharge reconstruction would give better results in basins with larger discharges (compared to the budget residual) such as the Amazon or the Orinoco basins (currently under study). When the original data sets are of poor quality, the integration process from SW+PF or CCM can only find the best consensus solution. However, in such a case, the reconstruction methodology could be improved by adding other constraints, such as the ocean and atmosphere water budgets or the energy budget. We will be investigating such an approach in the Mediterranean Basin.

We also demonstrated that the CCM can be used to reconstruct the water storage change prior to the beginning of the GRACE mission in 2002. The CCM calibrated over the period 2002–2010 was applied to correct P, E, and R between 1990 and 2002. Then ΔS was estimated using the water budget equation over this time period, as done in the gap filling method. The comparison with results from the *Pan et al.* [2012] study demonstrated the good performances of this model-independent method.

Compared with other studies dealing with the water budget closure at the basin scale, our methodology is based only on satellite observations and in situ runoff measurements. No reanalyses or models are used. As a consequence, the integrated data sets is model independent and can be used for model calibration or validation. Besides, the main advantages of the CCM are the following: (1) the integration method is very simple (linear regression), (2) the calibration reference is not a single data set but the consensus of multiple data sets of the four components (plus the closure constraint), (3) CCM is time independent (e.g., no trend or seasonal component in the budget residual), and (4) it is possible to downscale the integrated data sets by applying the CCM at each pixel.

Two main perspectives are envisaged. First, the integration methodology and the CCM will be evaluated over other basins. The basins should be large enough to be compatible with the GRACE spatial resolution and should be representative of the global environmental diversity. Possible candidates are the Amazon, the Congo, the Ob, or the Ganges basins. Second, we will check how the CCM varies from a basin to the other, for the same type of environments (hydroclimatic conditions) and across environments. Although it is very unlikely, if the closure correction model parameters are stable enough, it could be applied globally on the data sets, for all basins, even when any component of the hydrological cycle is missing. If the closure correction model parameters show significant environmental dependency, it could be parameterized to account for this dependency and applied globally. This environmental dependency would mean that the accuracy of each original satellite data set is different on different environments, which is a likely behavior. Provided that enough data are available to set up the CCM, this would mean that all the satellite data sets could be calibrated at the global scale, while getting closer to the water budget closure. One of the main limitations of performing a global-scale study is the availability of river discharge data. Such a study would benefit from the growing efforts of the community in developing algorithms to derive river discharge from remote sensing altimetry data. Such a study would fulfill one of the key objective of GEWEX Data and Evaluation Panel which is to provide consistent satellite-derived estimates of the water cycle components, first to analyze the past and current hydrological cycle under a changing climate, and second to evaluate the global hydrological models for a more accurate climate prediction. Efforts have been made in order to ensure such a closure of the water budget but mostly using global averages with yearly time resolutions. The work proposed using a global CCM would go beyond this by ensuring the correctness of the spatial and temporal patterns.

Acknowledgments
We would like to thank ESA (European Space Agency) and in particular Diego Fernandez for funding the project "Watchful: STSE Water Cycle Feasibility" (4000107122/12/I-NB) and Marcela Doubkova for managing this project. We would also like to thank colleagues Wolfgang Wagner, Richard Kidd, Wouter Dorigo, Patrick Matgen, and Paul Bates from the Watchful project for interesting related discussions. Data to support this article are available upon request to the corresponding author (simon.munier@gmail.com).

References

Adler, R. F., et al. (2003), The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present), *J. Hydrometeorol.*, *4*, 1147–1167.

Aires, F. (2014), Combining datasets of satellite retrieved products. Part I: Methodology and water budget closure, *J. Hydrometeorol.*, *15*, 1677–1691, doi:10.1175/JHM-D-13-0148.1.

Allen, R. G., L. S. Pereira, D. Raes, and M. Smith (1998), Crop evapotranspiration: Guidelines for computing crop water requirements. FAO irrigation and drainage paper 56. *FAO, Rome, 300*, 6541.

Azarderakhsh, M., W. B. Rossow, F. Papa, H. Norouzi, and R. Khanbilvardi (2011), Diagnosing water variations within the Amazon basin using satellite data, *J. Geophys. Res.*, *116*, D24107, doi:10.1029/2011JD015997.

Bjerklie, D. M., D. Moller, L. C. Smith, and S. L. Dingman (2005), Estimating discharge in rivers using remotely sensed hydraulic information, *J. Hydrol.*, *309*(1), 191–209.

Coumou, D., and S. Rahmstorf (2012), A decade of weather extremes, *Nat. Clim. Change*, *2*, 491–496.

Dai, A., K. E. Trenberth, and T. Qian (2004), A global dataset of Palmer Drought Severity Index for 1870–2002: Relationship with soil moisture and effects of surface warming, *J. Hydrometeorol.*, *5*, 1117–1130.

Dai, A., T. T. Qian, K. E. Trenberth, and J. D. Milliman (2009), Changes in continental freshwater discharge from 1948 to 2004, *J. Clim.*, *22*, 2773–2792.

Del Genio, A. D., A. A. Lacis, and R. A. Ruedy (1991), Simulations of the effect of a warmer climate on atmospheric humidity, *Nature*, *351*, 382–385.

Durand, M., J. Neal, E. Rodriguez, K. M. Andreadis, L. C. Smith, and Y. Yoon (2014), Estimating reach-averaged discharge for the River Severn from measurements of river water surface elevation and slope, *J. Hydrol.*, *511*, 92–104.

Ferguson, C. R., J. Sheffield, E. F. Wood, and H. Gao (2010), Quantifying uncertainty in a remote-sensing based estimate of evapotranspiration over continental USA, *Int. J. Remote Sens.*, *31*(14), 3821–3865.

Forootan, E., J. L. Awange, J. Kusche, B. Heck, and A. Eicker (2012), Independent patterns of water mass anomalies over Australia from satellite data and models, *Remote Sens. Environ.*, *124*, 427–443.

Gleason, C. J., and L. C. Smith (2014), Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry, *Proc. Nat. Acad. Sci.*, *111*(13), 4788–4791.

Gosset, M., J. Viarre, G. Quantin, and M. Alcoba (2013), Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks, *Q. J. R. Meteorol. Soc.*, *139*, 923–940.

Huffman, G. J., R. F. Adler, D. T. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker, and D. B. Wolff (2007), The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, *J. Hydrometeorol.*, *8*(1), 38–55.

Huntington, T. G. (2006), Evidence for intensification of the global water cycle: Review and synthesis, *J. Hydrol.*, *319*, 83–95.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie (2004), CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution, *J. Hydrometeorol.*, *5*, 487–503.

Jung, M., et al. (2010), Recent decline in the global land evapotranspiration trend due to limited moisture supply, *Nature*, *467*(7318), 951–954.

Kidd, C., P. Bauer, J. Turk, G. J. Huffman, R. Joyce, K. L. Hsu, and D. Braithwaite (2011), Intercomparison of high-resolution precipitation products over Northwest Europe, *J. Hydrometeorol.*, *13*(1), 67–83.

Klees, R., et al. (2008), A comparison of global and regional GRACE models for land hydrology, *Surv. Geophys.*, *29*(4-5), 335–359.

Labat, D., Y. Goddéris, J. L. Probst, and J. L. Guyot (2004), Evidence for global runoff increase related to climate warming, *Adv. Water Res.*, *27*, 631–642.

Landerer, F. W., and S. C. Swenson (2012), Accuracy of scaled GRACE terrestrial water storage estimates, *Water Resour. Res.*, *48*, W04531, doi:10.1029/2011WR011453.

Landerer, F. W., J. O. Dickey, and A. Güntner (2010), Terrestrial water budget of the Eurasian pan-Arctic from GRACE satellite measurements during 2003–2009, *J. Geophys. Res.*, *115*, D23115, doi:10.1029/2010JD014584.

Law, B. E., et al. (2006), Carbon fluxes across regions: Observational constraints at multiple scales, in *Scaling and Uncertainty Analysis in Ecology: Methods and Applications*, pp. 167–190, Springer, New York.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land-surface water and energy fluxes for general circulation models, *J. Geophys. Res.*, *99*, 14,415–14,428.

Liang, X., E. F. Wood, and D. P. Lettenmaier (1996), Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification, *Global Planet. Change*, *13*, 195–206.
Milliman, J. D., K. L. Farnsworth, P. D. Jones, K. H. Xu, and L. C. Smith (2008), Climatic and anthropogenic factors affecting river discharge to the global ocean, 1951–2000, *Global Planet. Change*, *62*, 187–194.
Milly, P. C. D., and K. A. Dunne (2001), Trends in evaporation and surface cooling in the Mississippi River Basin, *Geophys. Res. Lett.*, *28*, 1219–1222.
Miralles, D. G., T. R. H. Holmes, R. A. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman (2011), Global land-surface evaporation estimated from satellite-based observations, *Hydrol. Earth Syst. Sci.*, *15*(2), 453–469.
Mitchell, T. D., and P. D. Jones (2005), An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *Int. J. Climatol.*, *25*(6), 693–712.
Mu, Q., F. A. Heinsch, M. Zhao, and S. W. Running (2007), Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, *Remote Sens. Environ.*, *111*(4), 519–536.
Mu, Q., M. Zhao, and S. W. Running (2011), Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sens. Environ.*, *115*(8), 1781–1800.
Munier, S., H. Palanisamy, P. Masongrande, A. Cazenave, and E. F. Wood (2012), Global runoff anomalies over 1993–2009 estimated from coupled land-ocean-atmosphere water budgets and its relation with climate variability, *Hydrol. Earth Syst. Sci.*, *16*, 3647–3658.
Pan, M., and E. F. Wood (2006), Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter, *J. Hydrometeorol.*, *7*(3), 534–547.
Pan, M., and E. F. Wood (2010), Impact of accuracy, spatial availability, and revisit time of satellite-derived surface soil moisture in a multiscale ensemble data assimilation system, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, *3*(1), 49–56.
Pan, M., A. K. Sahoo, T. J. Troy, R. K. Vinukollu, J. Sheffield, and E. F. Wood (2012), Multisource estimation of long-term terrestrial water budget for major global river basins, *J. Clim.*, *25*(9), 3191–3206.
Prigent, C., F. Papa, F. Aires, C. Jimenez, W. B. Rossow, and E. Matthews (2012), Changes in land surface water dynamics since the 1990s and relation to population pressure, *Geophys. Res. Lett.*, *39*, L08403, doi:10.1029/2012GL051276.
Sahoo, A. K., M. Pan, T. J. Troy, R. K. Vinukollu, J. Sheffield, and E. F. Wood (2011), Reconciling the global terrestrial water budget using satellite remote sensing, *Remote Sens. Environ.*, *115*(8), 1850–1865.
Sapiano, M. R. P., and P. A. Arkin (2009), An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data, *J. Hydrometeorol.*, *10*, 149–166.
Scheel, M. L. M., M. Rohrer, C. Huggel, D. Santos Villar, E. Silvestre, and G. J. Huffman (2011), Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution, *Hydrol. Earth Syst. Sci.*, *15*(8), 2649–2663.
Schneider, U., T. Fuchs, A. Meyer-Christoffer, and B. Rudolf (2008), Global precipitation analysis products of the GPCC, Global Precipitation Climatology Centre (GPCC). *DWD, Internet Publication*. [Available at http://www.dwd.de, 1-12, 2008.]
Sheffield, J., C. R. Ferguson, T. J. Troy, E. F. Wood, and M. F. McCabe (2009), Closing the terrestrial water budget from satellite remote sensing, *Geophys. Res. Lett.*, *36*, L07403, doi:10.1029/2009GL037338.
Su, F., Y. Hong, and D. P. Lettenmaier (2008), Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin, *J. Hydrometeorol.*, *9*(4), 622–640.
Syed, T. H., J. S. Famiglietti, J. Chen, M. Rodell, S. I. Seneviratne, P. Viterbo, and C. R. Wilson (2005), Total basin discharge for the Amazon and Mississippi River basins from GRACE and a land-atmosphere water balance, *Geophys. Res. Lett.*, *32*, L24404, doi:10.1029/2005GL024851.
Syed, T. H., J. S. Famiglietti, M. Rodell, J. Chen, and C. R. Wilson (2008), Analysis of terrestrial water storage changes from GRACE and GLDAS, *Water Resour. Res.*, *44*, W02433, doi:10.1029/2006WR005779.
Tapley, B. D., S. Bettadpur, M. M. Watkins, and C. Reigber (2004), The gravity recovery and climate experiment: Mission overview and early results, *Geophys. Res. Lett.*, *31*, L09607, doi:10.1029/2004GL019920.
Thiemig, V., R. Rojas, M. Zambrano-Bigiarini, V. Levizzani, and A. de Roo (2012), Validation of satellite-based precipitation products over sparsely gauged African River Basins, *J. Hydrometeorol.*, *13*, 1760–1783.
Tian, Y., and C. D. Peters-Lidard (2010), A global map of uncertainties in satellite-based precipitation measurements, *Geophys. Res. Lett.*, *37*, L24407, doi:10.1029/2010GL046008.
Trenberth, K. E. (1999), Conceptual framework for changes of extremes of the hydrological cycle with climate change, *Clim. Change*, *42*, 327–339.
Turk, J. T., G. V. Mostovoy, and V. Anantharaj (2010), The NRL-blend high resolution precipitation product and its application to land surface hydrology, in *Satellite Rainfall Applications for Surface Hydrology*, edited by M. Gebremichael, and F. Hossain, pp. 85–104, Springer, Netherlands.
Villarini, G., W. F. Krajewski, and J. A. Smith (2009), New paradigm for statistical validation of satellite precipitation estimates: Application to a large sample of the TMPA 0.25° 3-hourly estimates over Oklahoma, *J. Geophys. Res.*, *114*, D12106, doi:10.1029/2008JD011475.
Werth, S., and A. Güntner (2010), Calibration analysis for water storage variability of the global hydrological model WGHM, *Hydrol. Earth Syst. Sci.*, *14*(1), 59–78.
Willmott, C. J., and K. Matsuura (2010), Terrestrial air temperature and precipitation: Monthly and annual time series (1900–2008), V2.01. Center of Climatic Research, Univ. of Delaware.
Wu, H., R. F. Adler, Y. Tian, G. J. Huffman, H. Li, and J. Wang (2014), Real-time global flood estimation using satellite-based precipitation and a coupled land surface and routing model, *Water Resour. Res.*, *50*, 2693–2717, doi:10.1002/2013WR014710.
Zaitchik, B. F., M. Rodell, and R. H. Reichle (2008), Assimilation of GRACE terrestrial water storage data into a Land Surface Model: Results for the Mississippi River basin, *J. Hydrometeorol.*, *9*(3), 535–548.
Zhang, K., J. S. Kimball, R. R. Nemani, and S. W. Running (2010), A continuous satellite-derived global record of land surface evapotranspiration from 1983 to 2006, *Water Resour. Res.*, *46*, W09522, doi:10.1029/2009WR008800.

